

FIGURE 1

The Universal Code Equivalent of  
the Mitochondrial 1-Sce I Gene.

AAAAAATAAAATCAT ATG AAA AAT ATT AAA AAA AAT CAA GTA ATG AAT CTC GGT GGT AAT TCT  
 M K N I K K H Q Y H H L S P H S  
 AAA TTA TTA AAA GAA TAT AAA TCA CAA TTA ATT GAA TTA AAT ATT GAA CAA TTT GAA GAA  
 K L L K E Y K S Q L I E L H I E Q F E A  
 GGT ATT GGT TTA ATT TTA GGA GAT GGT TAT ATT GGT AGT GGT GAT GAA GGT AAA ACT TAT  
 G I G L I L G D A Y I R S R D E S K T Y  
 TGT ATG CAA TTT CAC TCC AAA AAT AAG GCA TAC ATG GAT CAT GTA TGT TTA TTA TAT GAT  
 C H Q F E W K H K A Y H D H Y C L L Y D  
 CAA TGG GTA TTA TCA CCT CCT CAT AAA AAA GAA AGA GTT AAT CAT TTA GGT AAT TTA GTA  
 Q W Y L S P P H K K E R V H H L G H L Y  
 ATT ACC TGG GGA GGT CAA ACT TTT AAA CAT CAA GGT TTT AAT AAA TTA GGT AAC TTA TTT  
 I T W G A Q T F K H Q A F H K L A H L F  
 ATT GTA AAT AAT AAA AAA CTT ATT CCT AAT AAT TTA GTT GAA AAT TAT TTA ACA GGT ATG  
 I Y H H K K L I P H H L Y E H Y L T P H  
 AGT CTG GCA TAT TGG TTT ATG GAT GAT GGA GGT AAA TGG GAT TAT AAT AAA AAT TCT GGT  
 S L A Y W F H D D Q G K W D Y H K H S L  
 AAT AAA AGT ATT GTA TTA AAT ACA CAA AGT TTT ACT TTT CAA GAA GTA GAA TAT TTA CTT  
 H K S I V L H T Q S F T F E E V C Y L V  
 AAA GGT TTA AGA AAT AAA TTT CAA TTA AAT TGT TAT GTT AAA ATT AAT AAA AAT AAA CCA  
 K G L R H K F Q L H C Y Y K I H K H K P  
 ATT ATT TAT ATT GAT TCT ATG AGT TAT CTG ATT TTT TAT AAT TTA ATT AAA CCT TAT TTA  
 I I Y I D S H S Y L I F Y H I I K P Y L  
 ATT CCT CAA ATG ATG TAT AAA CTG CCT AAT ACT ATT TCA TCC GAA ACT TTT TTA AAA TAA  
 I P Q H H Y K L P H T I S S E T F L K

6610220" 06111260

FIGURE 2

The synthetic I-Sce I gene

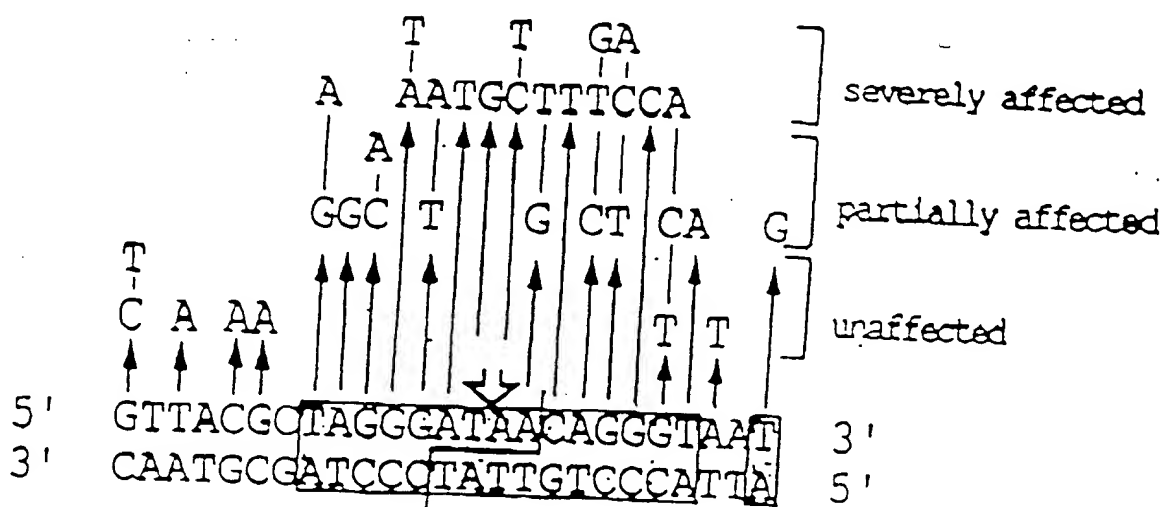
Bam HI  
|  
CCGGATCCATG CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GGT CCG AAC TGT  
M H M K N I K K N Q Y M N L G P N S  
AAA CTG CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA GCA  
K L L K E Y K S Q L I E L N I E C F E A  
1. GGT ATC GGT CTG ATC CTG GGT GAT GGT TAC ATC CGT TCT CGT GAT GAA GGT AAA ACC TAC  
G I G L I L G D A Y I R S R D E G K T Y  
TGT ATG CAG TTC CAG TGG AAA AAC AAA GCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT  
C M Q F E W K N K A Y M D H V C L L Y D  
CAG TGG GTA CTG TCC CCG CCG CAC AAA AAA CAA CGT GTT AAC CAC CTG GGT AAC CTG GTA  
Q W V L S P P H K K E R Y N H L G N L Y  
ATC ACC TGG GGC GGC CAG ACT TTC AAA CAC CAA GCT TTC AAC AAA CTG GCT AAC CTG TTC  
I T W G A Q T F K M Q A F N K L A N L F  
ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA AAC TAC CTG ACC CCG ATG  
I Y N N K K T I P N N L Y E N Y L T P M  
2. TCT CTG GCA TAC TGG TTC ATG GAT GAT GGT GGT AAA TGG GAT TAC AAC AAA AAC TCT ACC  
S L A Y W F M D D G K W D Y N K N S T  
AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA GAA GTA GAA TAC CTG GTT  
N K S I V L N T Q S F T F E E V E Y L Y  
AAC GGT CTG CGT AAC AAA TTC CAA CTG AAC TGT TAC GTA AAA ATC AAC AAA AAC AAA CCG  
K G L R N K F Q L N C Y Y K I H K N K P  
ATC ATC TAC ATC GAT TCT ATG TCT TAC CTG ATC TTC TAC AAC CTG ATC AAA CCG TAC CTG  
I I Y I D S M S Y L I F Y N L I K P Y L  
ATC CCG CAG ATG ATG TAC AAA CTG CCG AAC ACT ATC TCC TCC GAA ACT TTC CTG AAA TAA  
I P Q M M Y K L P N T I S S E T F L K  
TAAGTCGACTGCAGGATCCGGTAAGTAAGTAA  
| | |  
Sall PstI BamHI

1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.

660000 DE 111000

FIGURE 3

66000" 06711200





# VARIATIONS AROUND THE 1-Scp 1 SEQUENCE

	-2	-1	1	2	3	4	5	6	7	8	9	10							
	M	H	V	Y	V	I	C	X	V	I	V	N	V	L	C	P	V	S	
			20									30							
K	L	L	K	E	Y	K	S	D	L	I	E	L	N	I	E	C	F	E	A
			40									50							
G	I	S	L	I	L	A	A	Y	I	R	S	R	D	E	G	K			
			60									70							
C	H	D	F	E	H	K	H	K	A	Y	H	D	H	V	C	L			
			80									90							
Q	V	Y	L	S	P	P	H	K	K	E	R	Y	N	H	L	G	H		
			100									110							
I	T	V	G	A	D	T	F	K	H	D	A	F	N	K	L	A	N		
			120									130							
I	V	N	H	K	K	I	I	P	N	N	L	V	E	H	Y	L	T	D	
			140									150							
D	L	A	Y	V	F	H	D	D	D	D	K	V	D	Y	H	K	H	S	
			160									170							
H	K	S	I	V	L	N	T	D	S	F	T	F	E	E	V	E	Y		
			180									190							
K	G	L	R	N	K	F	D	L	N	C	Y	V	K	I	H	K	N	K	D
			200									210							
I	I	Y	I	D	S	H	S	Y	L	I	F	Y	H	L	I	K	P	Y	
			220									230							
I	P	Q	H	H	Y	K	L	P	H	T	I	S	S	E	T	F	L	K	

## Positions that can be changed without affecting enzyme activity (demonstrated)

positions -1 and -3 are not natural. The two amino acids are added due to cloning strategies

positions 1 to 10: can be deleted

position 36: G is tolerated

position 40: M or Y are tolerated

position 41: S or H are tolerated

position 43: A is tolerated

position 46: Y or N are tolerated

position 91: A is tolerated

positions 153 and 154: L are tolerated

position 222: A and S are tolerated

## Changes that affect enzyme activity (demonstrated)

position 19: L to S

position 38: I to S or N

position 39: G to D or R

position 40: L to Q

position 42: L to R

position 44: D to E, G or H

position 45: A to K or D

position 46: Y to D

position 47: I to R or Y

position 50: L to S

position 144: D to E

position 145: D to E

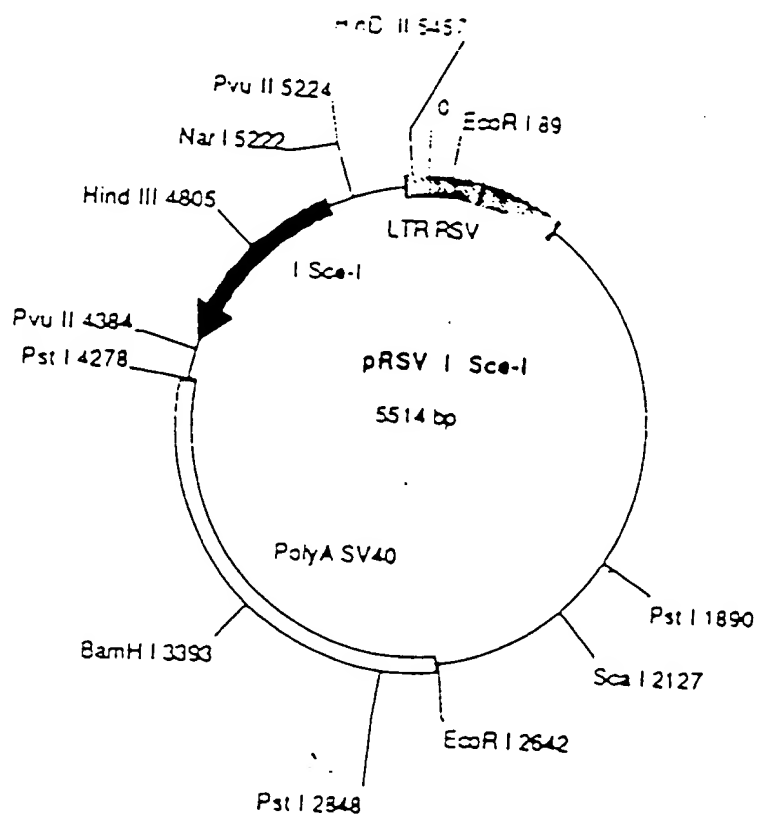
660300"0ET000

## Group I Intron Encoded Endonucleases and Related Endonucleases

Endonuclease		Recognition Sequence	CLAVY/CX SITS	▼ Intron site
Two Dodecapptide Family (for 4 bp cutters)	I-Sce I ( <i>Saccharomyces mitochondria</i> )	CGCTAGGGGATAA CAGGGTAAATATAGC GCGATCCCTATATATATATATATATATAT		
	I-Sce IV ( <i>Saccharomyces mitochondria</i> )	TTCTCATGATCA AGCTTAAATCCATAG AAGAGTAC TAATCAGATAGGATAGC		
	I-Sce II ( <i>Saccharomyces mitochondria</i> )	CTTTGGTCA CCAGAAAGTATATATAT GAATCCG GATAGGATATATATATATAT		
	I-Ceu I ( <i>Chlamydomonas chloroplast</i> )	TAACGGTCC TAAGCTTACCGAAATCTA ATTGCCAG GATTCGATCGGCTTAAAT		
	I-Hpo I ( <i>Physarum nucleus</i> )	TGACCTCTC TAAAGGATAGC AAATCGC ACTGAGCAG ATATATATATATATATAT		
	I-Sce III ( <i>Saccharomyces mitochondria</i> )	CGAGGTTTTT GTAACTATTTATTTACG CCTCCAAAGUATTGATTAATATATAT		
	I-Cre I ( <i>Chlamydomonas chloroplast</i> )	GGGTTCAAAACCT CGTGGAGACAGTTT CCCAGTTTTT CAGACATCTCTCTATA		
	Endo Sec I (RF3) ( <i>Saccharomyces mitochondria</i> ) (Not in vivo)	GATGCTG AGGCATAGGCTTATATAT GTACGACAT TCCCTATCGGAACCAATA		
	HO ( <i>Saccharomyces nucleus</i> ) (Not in vivo)	CTTCCGCG ACAGTATATATTTTATAA TAAAGGCT ATATATATATATATATAT		
	I-Cem I ( <i>Chlamydomonas mitochondria</i> ) (Putative endonuclease)	ACCATGGGGT CAAATATCTTTCTTGGG TGGTACCCCGATTTACAGAAACACCC		
	I-Pan I ( <i>Podosporea mitochondria</i> ) (Putative endonuclease)	GGCCCTGAATG ATATATATATATATAT CAGCCACTTACTATATATATATATAT		
Other Structural Families	(Bacteriophage T4)			
	I-Tev I	GAAAG CCCTCAGTACATTTTTCTTGGCTCTACCGTTTAA GTTCGAGTCACTATCAAAAAGAACCCAGATGGCAATTA		
	I-Tev II	CAAGCTTATGATATGAAGTCAACACCGT TAT GTTCGAA ACTCACTTCACTTGTG CAATTA		
	I-Tev III	GCTATTCGTTCTT TATGTATCTTTTGGCTTACCTTTAA CCATTAAGCAAAAT CATAGAAAACCCACATGCAAT		

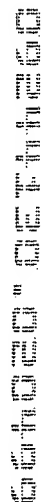


FIGURE 8



603330-000000

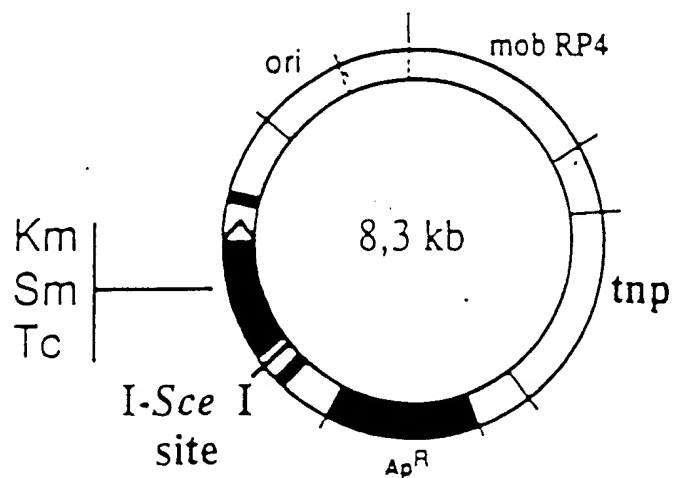






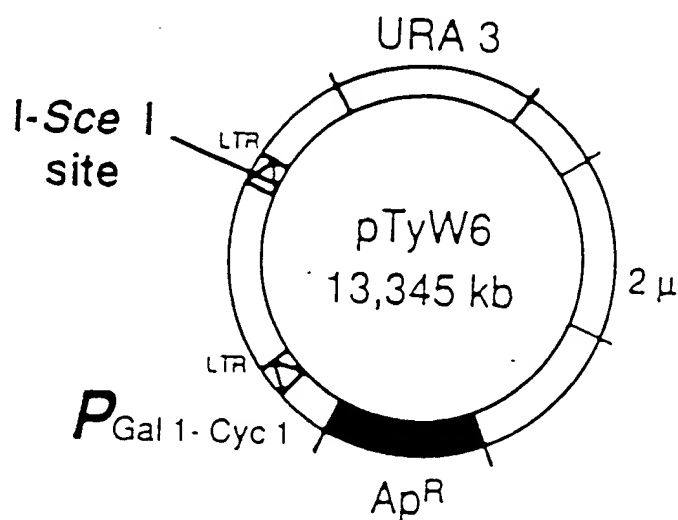
[illegible][illegible]

Names : pT $\omega$ Sm  
 pT $\omega$ Km  
 pT $\omega$ Tc

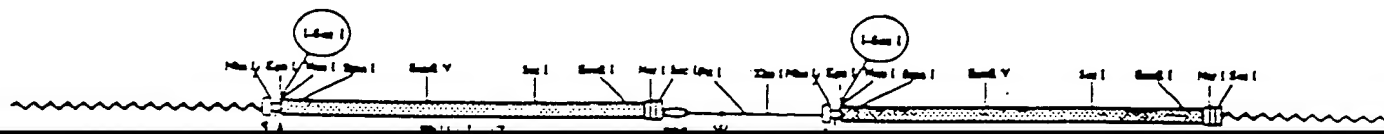
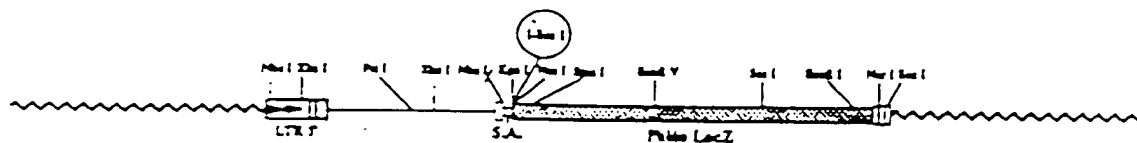
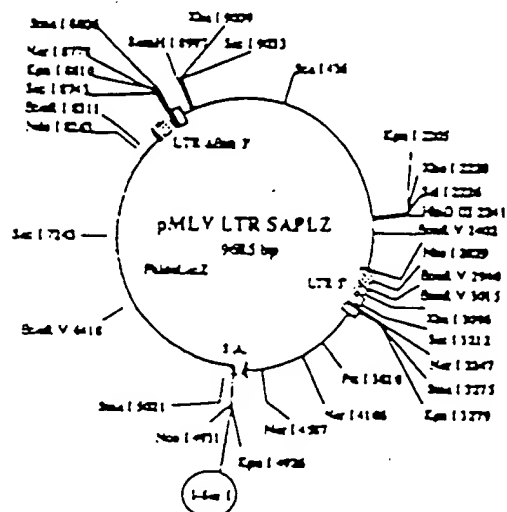


Construction : pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-SceI] in NotI unique site

FIGURE 12



Construction : pD 123 , from J. D. Boeke  
with insertion of a linker [ I-SceI - NotI ] in BamHI



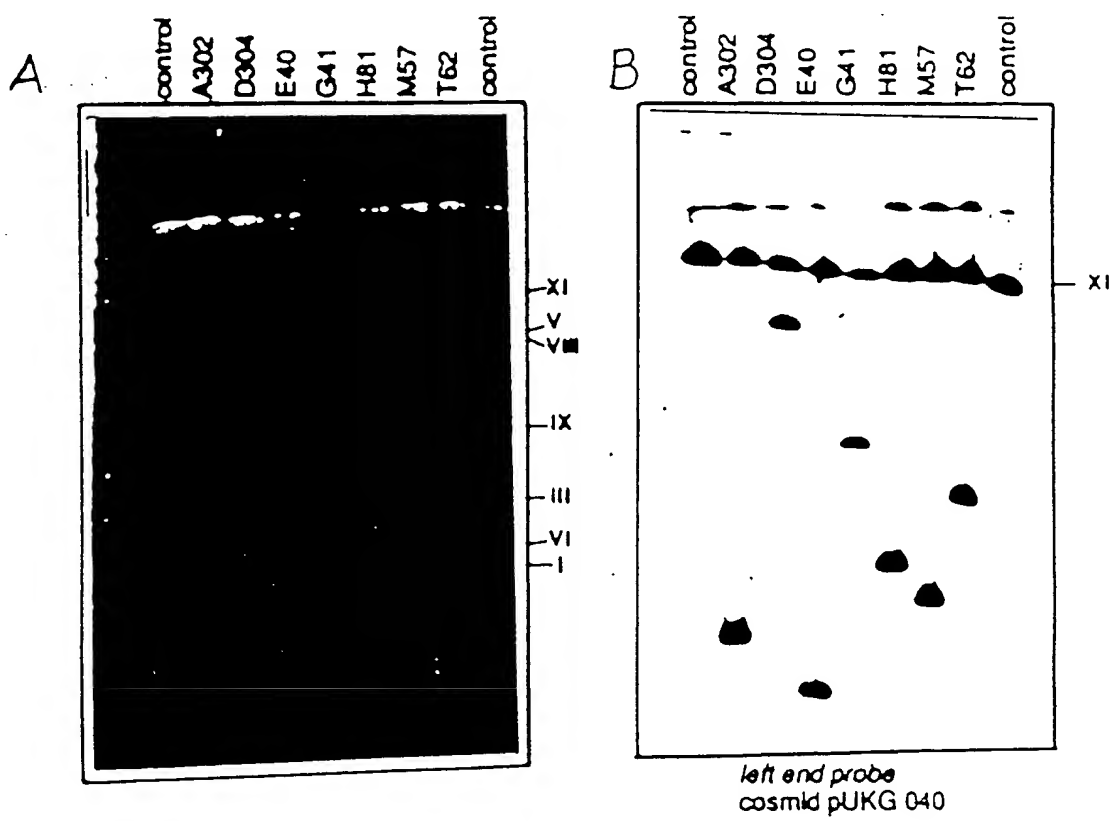


FIGURE 15

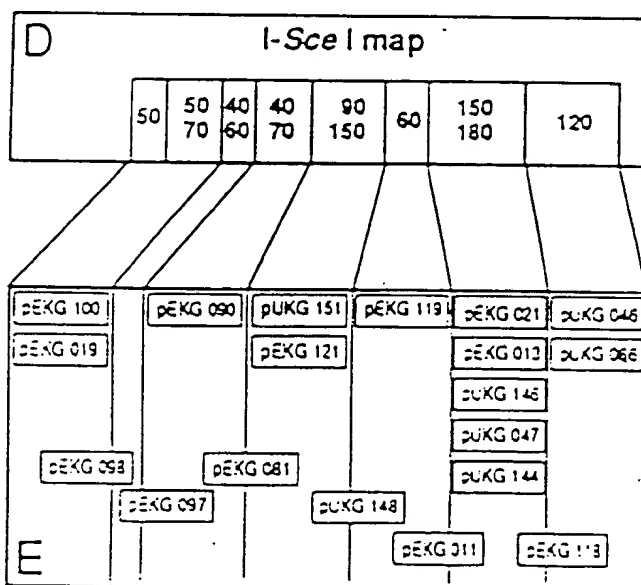
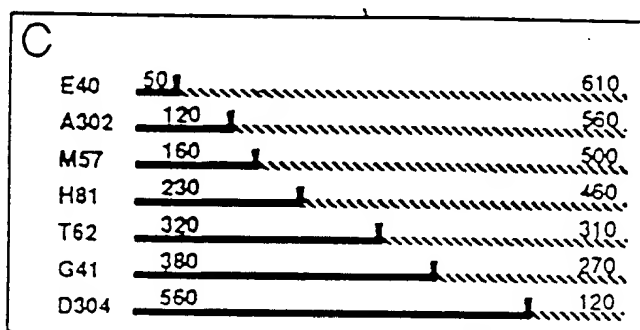
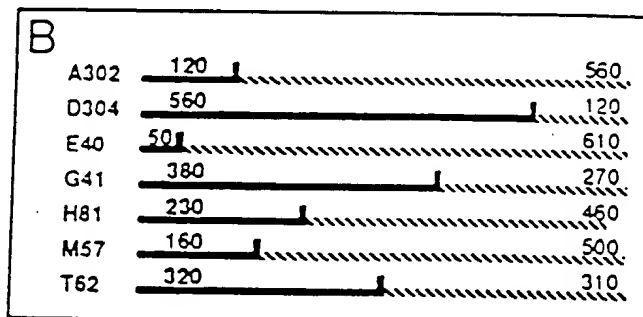
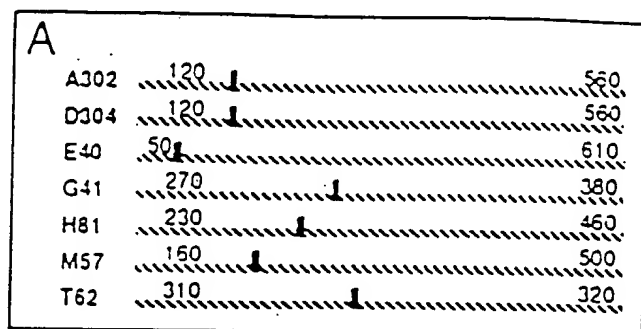




FIGURE 16

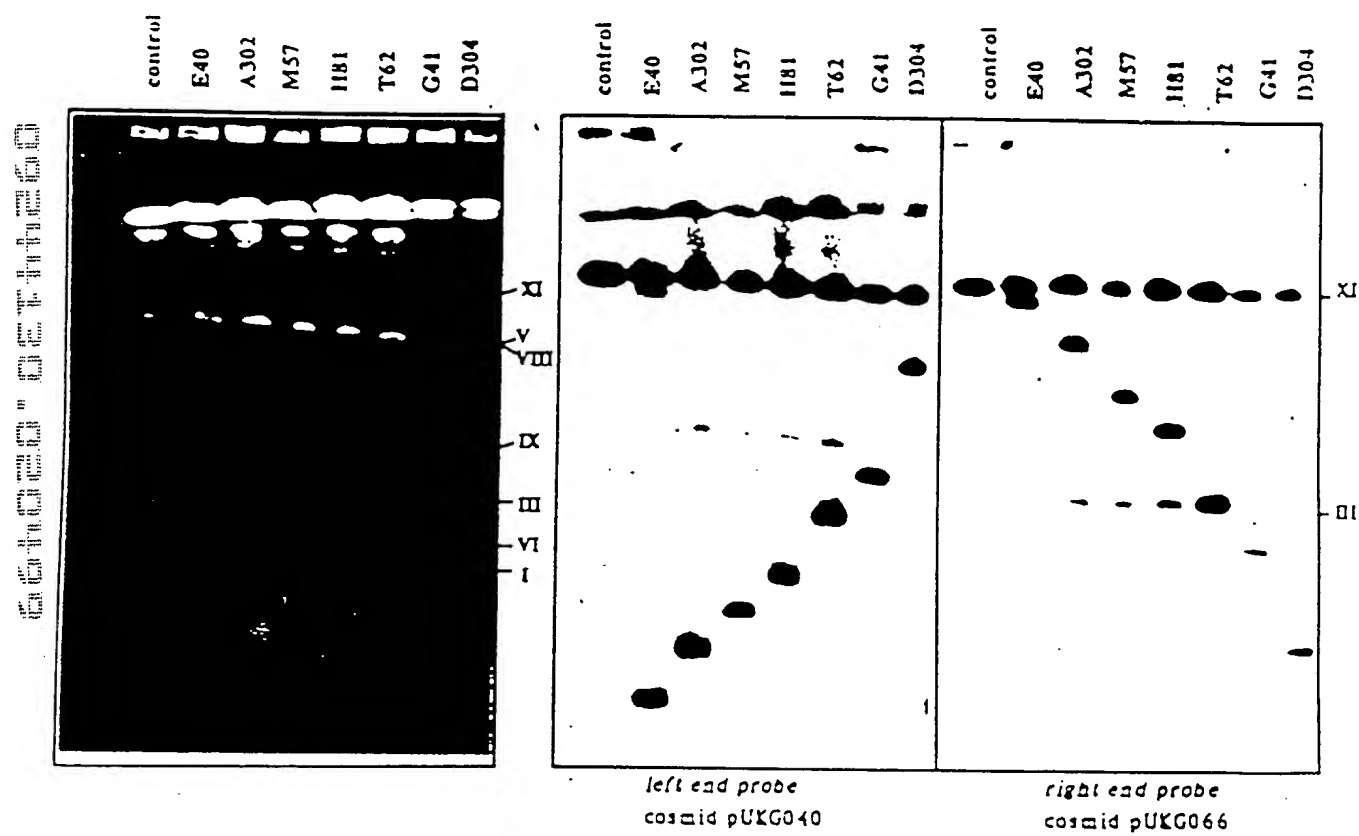
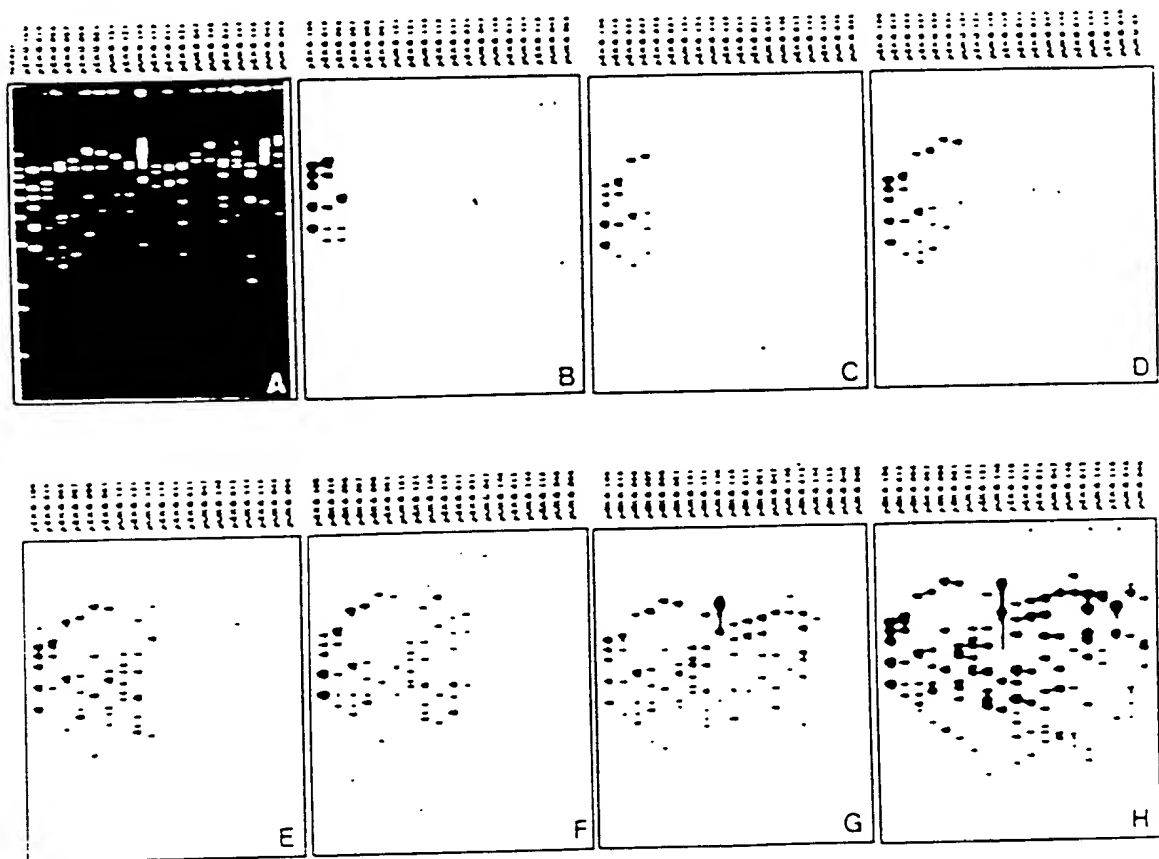
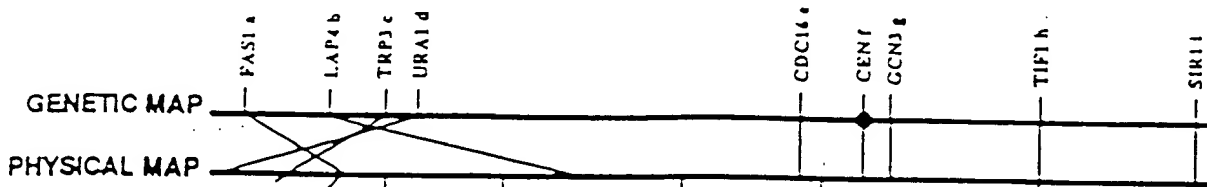


FIGURE 17





TEL	64 kb	64-78 kb	40-48 kb	44-78 kb	84-120 kb	80 kb	150-198 kb	128 kb	TEL
	pUKG 145 d	pEKG 025	pEKG 107	pEKG 006	pUKG 128 b	pEKG 002	pEKG 021	pUKG 064 h	
	pUKG 046 ad	pEKG 002	pEKG 007	pUKG 130	pEKG 002 b	pUKG 072	pUKG 128 i	pUKG 130	
	pEKG 012 ad	pUKG 129	pEKG 008	pEKG 024	pUKG 141 b	pUKG 148	pUKG 064 i	pEKG 104	
	pUKG 033 ad	pUKG 074	pEKG 007	pUKG 005	pUKG 151 b	pEKG 016	pUKG 047 i	pEKG 115	
	pUKG 077 ad	pEKG 010			pUKG 078 b	pUKG 001	pUKG 037 g	pEKG 116	
	pEKG 104 ad	pUKG 150			pEKG 006 b	pUKG 002	pEKG 006 g	pUKG 140	
		pUKG 136 a			pUKG 137 b	pEKG 004	pUKG 125 f	pEKG 120	
					pUKG 188 b	pUKG 064	pUKG 030 g	pEKG 064	
					pEKG 028	pEKG 008	pEKG 003 g	pEKG 100	
					pEKG 114	pEKG 110	pUKG 034 g	pUKG 152	
					pEKG 121	pUKG 003	pEKG 015 g	pUKG 046	
					pEKG 112	pUKG 022	pUKG 076 g	pUKG 052	
					pEKG 113	pEKG 110	pEKG 013 g	pUKG 131	
					pUKG 150	pEKG 027	pUKG 045	pUKG 136 i	
					pEKG 105	pEKG 014	pEKG 006	pEKG 101 i	
					pEKG 122	pEKG 008	pUKG 073	pEKG 007 i	
					pUKG 070	pUKG 007		pEKG 006 i	
					pEKG 020			pUKG 066 i	
								pUKG 123 i	
	pUKG 134 c								
	pEKG 023 c								
	pEKG 008 c								
	pEKG 001 c								
	pUKG 008 c	pEKG 065 a	pUKG 154	pEKG 001	pEKG 080	pEKG 011 e	pUKG 035 h		
	pEKG 017 c	pUKG 132 a	pUKG 057	pEKG 001	pUKG 135	pUKG 036 e	pUKG 124 h		
	pUKG 055 c	pEKG 008 a	pUKG 037	pEKG 108	pUKG 140	pEKG 111 e	pEKG 008 h		
	pUKG 051 c	pUKG 157 a	pUKG 053	pUKG 155	pUKG 133	pUKG 147 e	pEKG 106 h		
	pEKG 084 c	pUKG 075 a	pUKG 058	pUKG 040		pEKG 010 e	pEKG 103 h		
	pUKG 029 c	pEKG 003 a		pUKG 030		pUKG 041	pEKG 003 h		
	pUKG 143 c	pUKG 142 a		pUKG 048			pUKG 144 h		

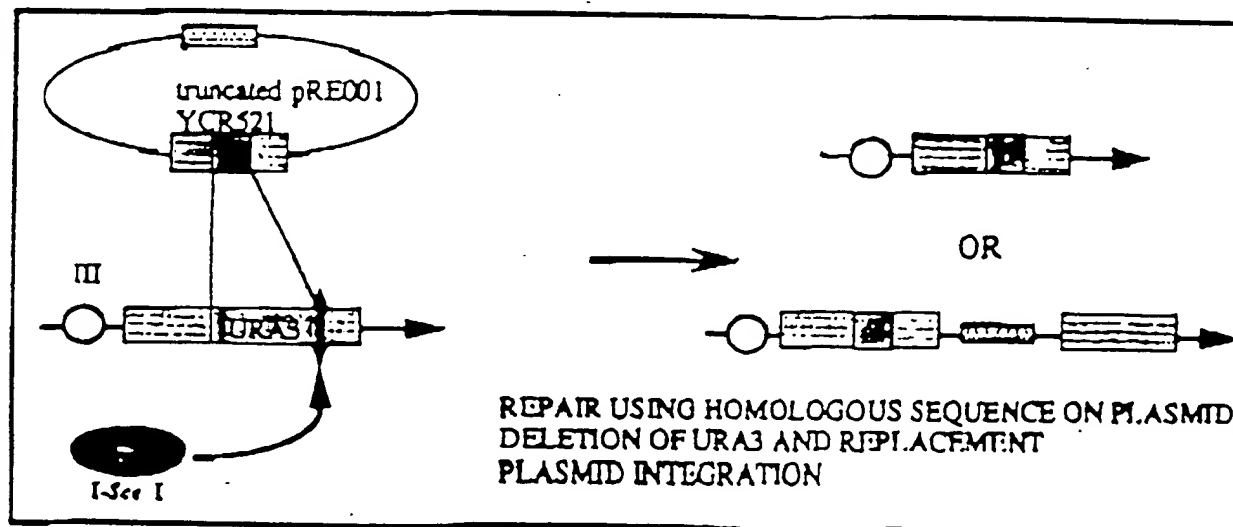
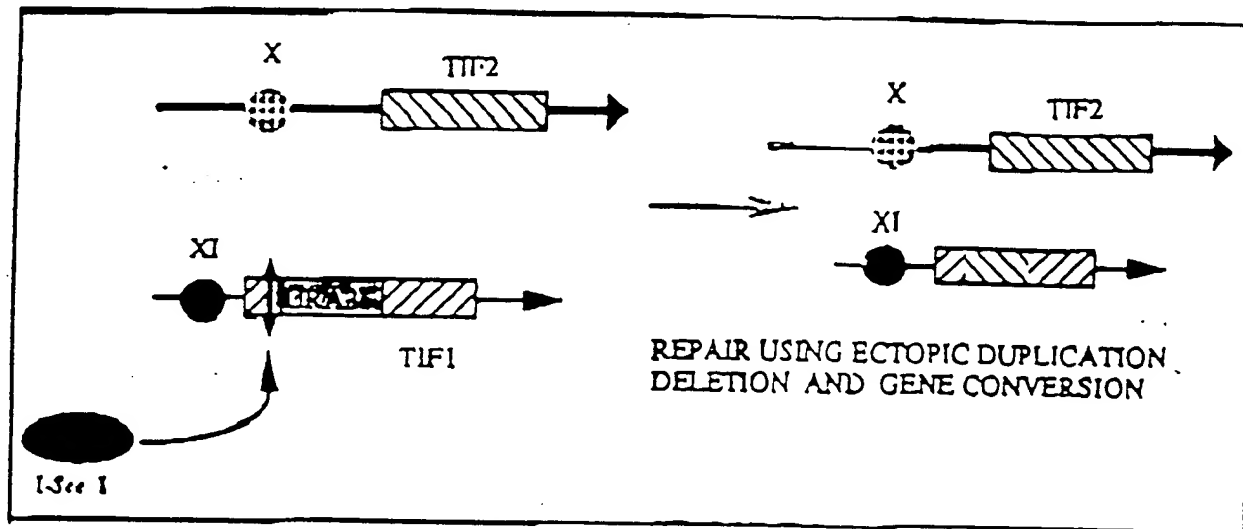


Figure 20

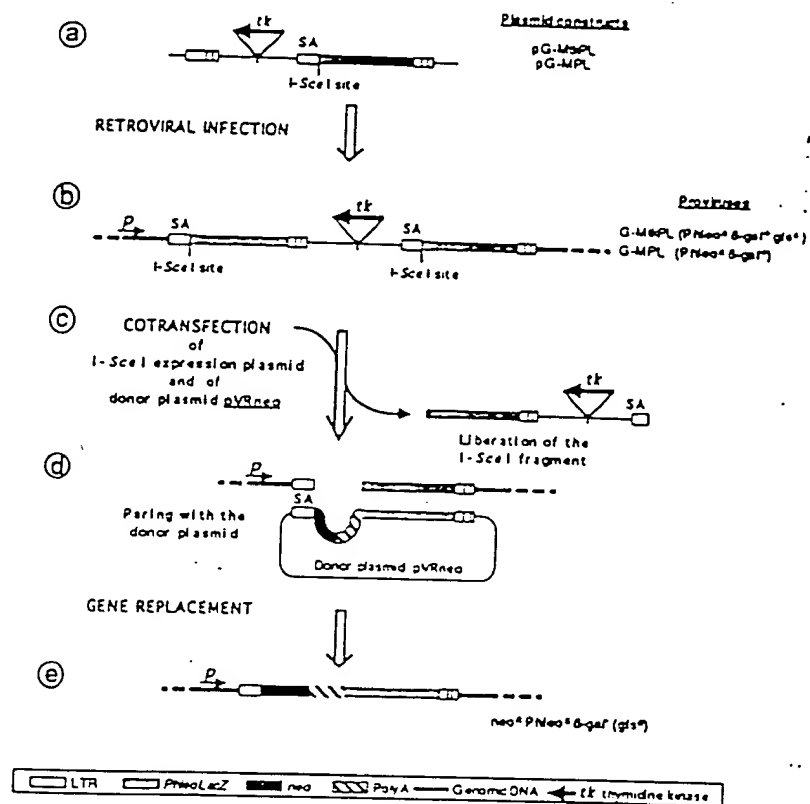
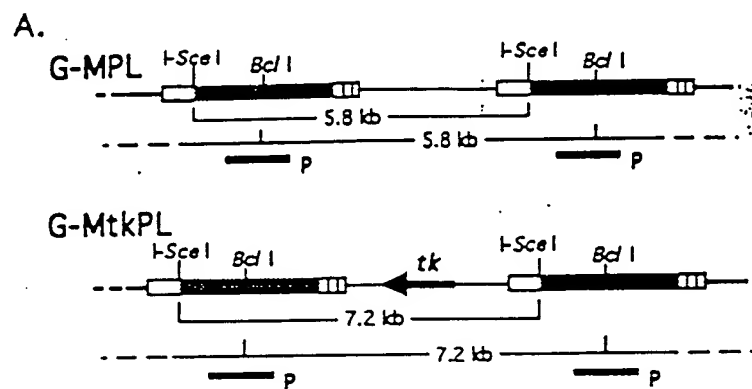


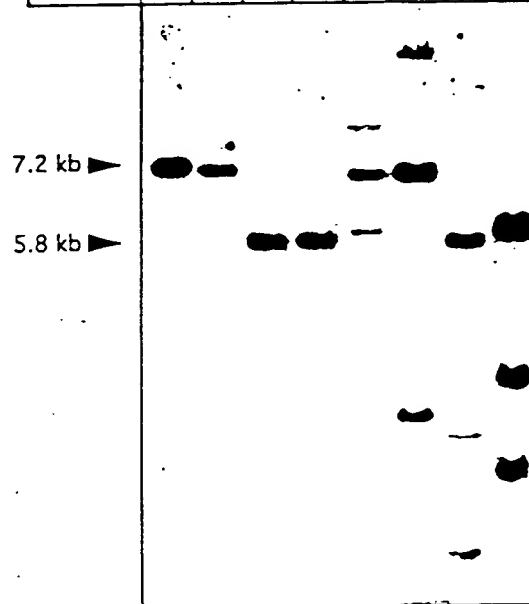


Figure 22



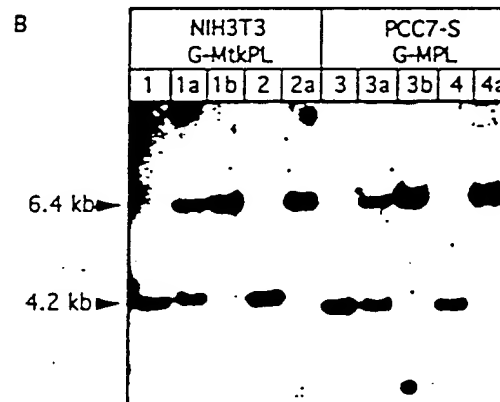
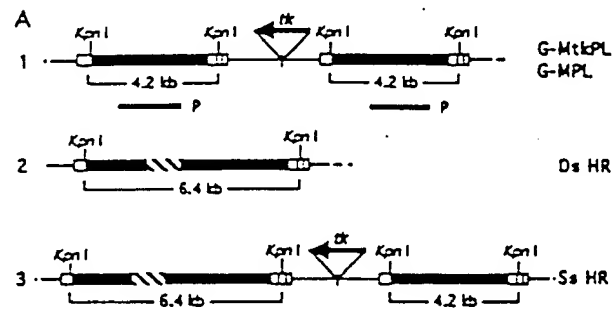
B.

ENZYME	I-SceI				BclI			
	NIH3T3		PCC7-S		NIH3T3		PCC7-S	
CELL LINE	G-MtkPL		G-MPL		G-MtkPL		G-MPL	
PROVIRUS	G-MtkPL		G-MPL		G-MtkPL		G-MPL	
CLONE	1	2	3	4	1	2	3	4



LTR
 PhleoLacZ
 tk thymidine kinase

Figure 23



LTR
PhleoLucZ
neo
PolyA
 Genomic DNA
 ← tk thymidine kinase



Figure 24

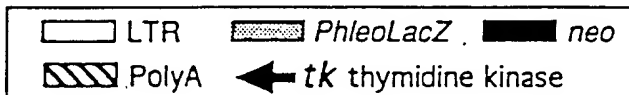
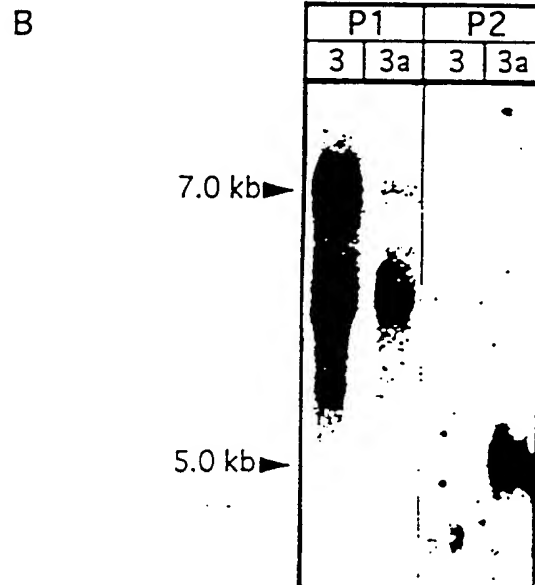
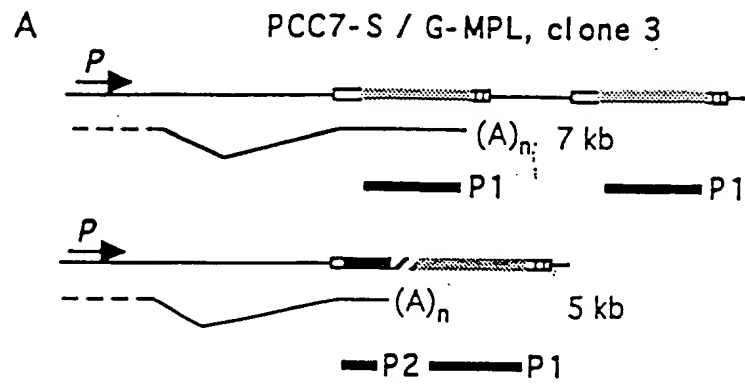
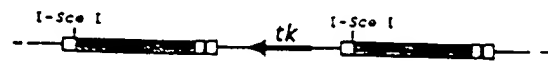


Figure 25

a. Chromosomal DNA containing provirus

Phenotypes

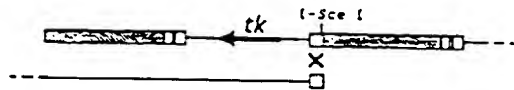


[Phleo<sup>R</sup>, Glu<sup>R</sup>, Δ-Gal<sup>+</sup>]

Transfection by  
I-Sce I endonuclease  
expression vector

b. Intra-chromosomal recombinations events

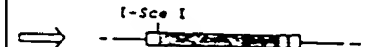
1. The left I-Sce I site is cut.  
Pairing and recombination



2. The right I-Sce I site is cut.  
Pairing and recombination



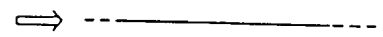
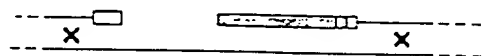
3. Both I-Sce I sites are cut.  
Religation by end-joining



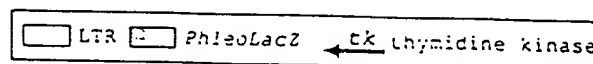
[Phleo<sup>R</sup>, Glu<sup>R</sup>, Δ-Gal<sup>+</sup>]

c. Inter-chromosomal recombination event

Both I-Sce I sites are cut. Gap repair using  
intact chromosome sequences

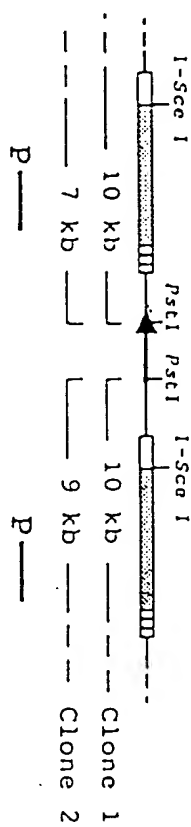


[Phleo<sup>R</sup>, Glu<sup>R</sup>, Δ-Gal<sup>-</sup>]



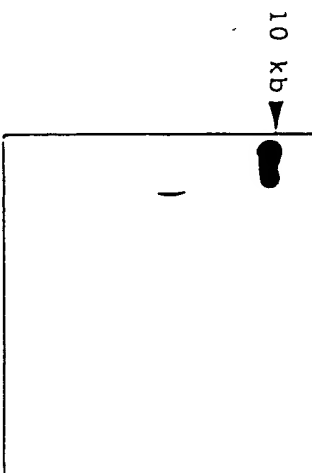
~~~~~~

a. Parental DNA, G-MtkPL



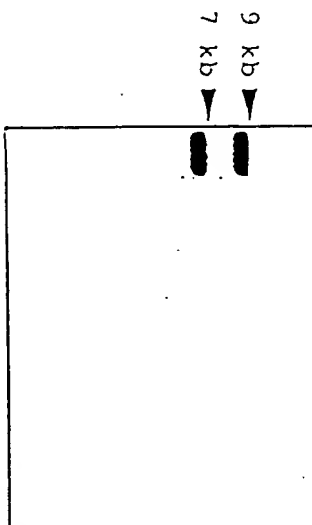
४.

| Pct I  |   |     |     |     |     |
|--------|---|-----|-----|-----|-----|
| ENZYME |   |     |     |     |     |
| CLONE  | 1 | 1.1 | 1.2 | 1.3 | 1.4 |
|        |   |     |     |     | 1.5 |



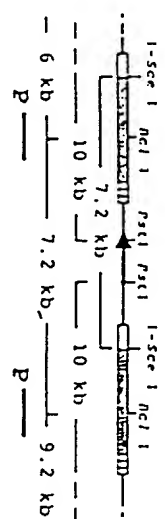
Q.

| ENZYME | Pct I |     |     |     |     |     |     |
|--------|-------|-----|-----|-----|-----|-----|-----|
| CLONE  | 2     | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 2.6 |



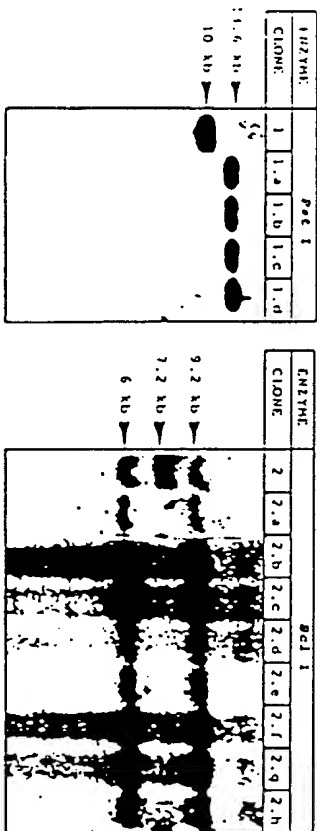
# C O N T E N T S

2000

[illegible]

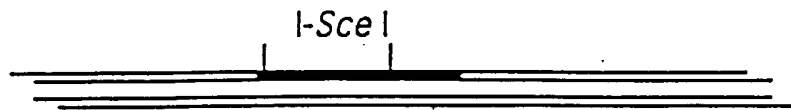
Q

|        |   | SC1 1 |     |     |     |     |     |     |     |
|--------|---|-------|-----|-----|-----|-----|-----|-----|-----|
| ENZYME |   |       |     |     |     |     |     |     |     |
| CLONE  | 2 | 2.a   | 2.b | 2.c | 2.d | 2.e | 2.f | 2.g | 2.h |

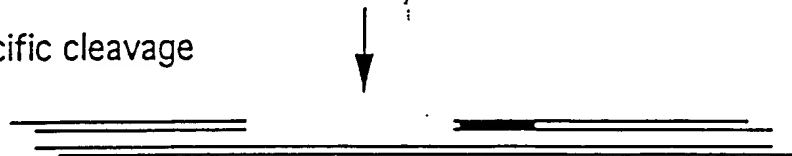
[illegible]

# LOSS OF HETEROZYGOSITY

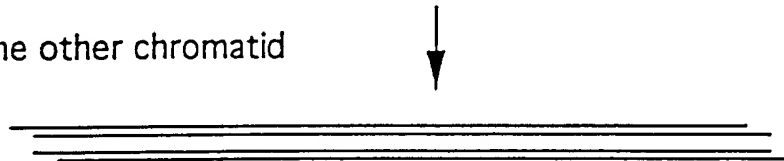
Integration of artificial site or presence of specific site



Expression of I-Sce I and specific cleavage



Repair of the DSB with the other chromatid

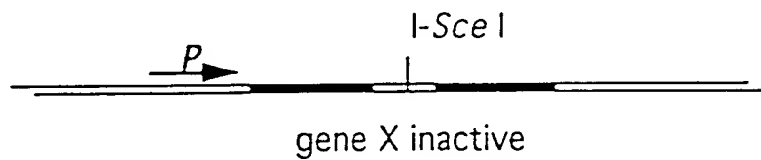


6610220 DET 11260

Figure 29

## CONDITIONAL ACTIVATION (Tandem repeat)

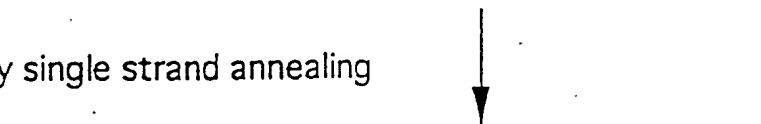
Integration of artificial site between tandem repeats



Expression of I-Sce I and specific cleavage



Repair of the DSB by single strand annealing

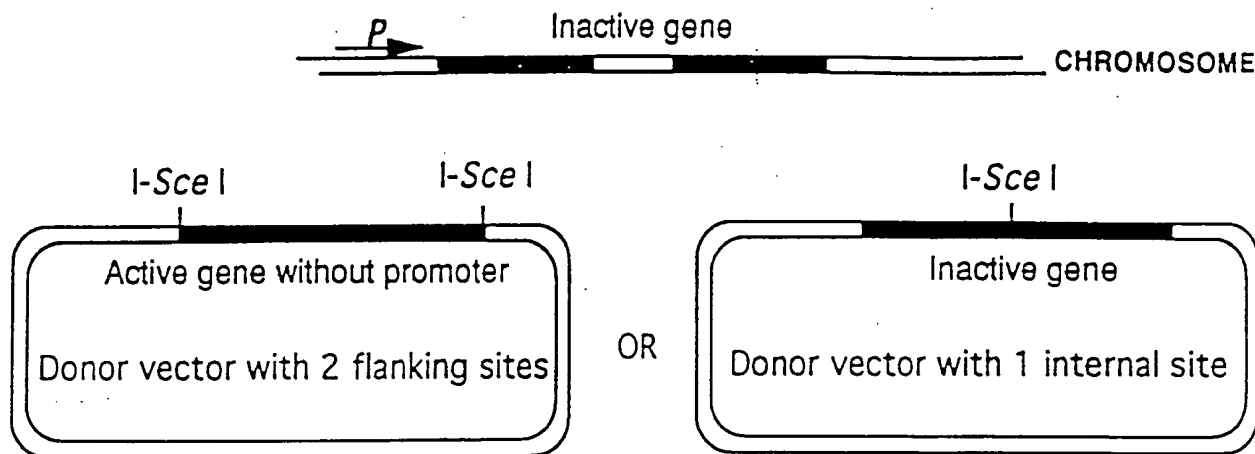


gene X active

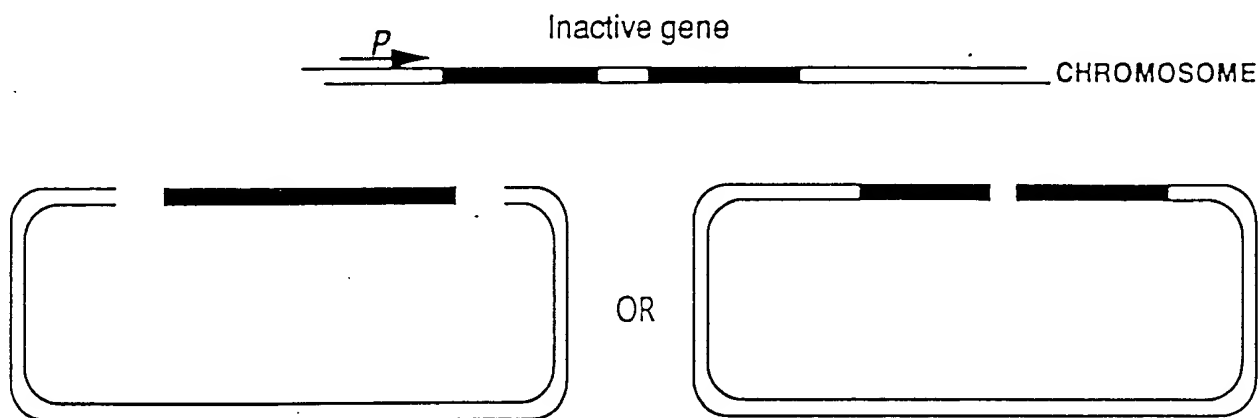
# ONE STEP REARRANGEMENT

Figure 30

Integration of artificial site or presence of specific site



Expression of I-Sce I enzyme  
and  
specific cleavage of the donor plasmid



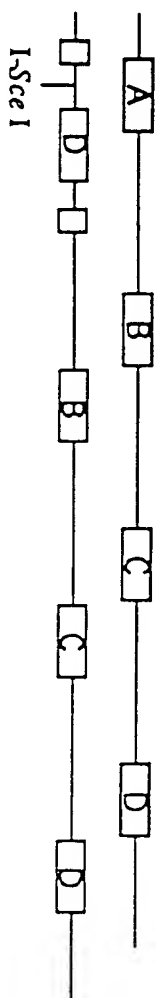
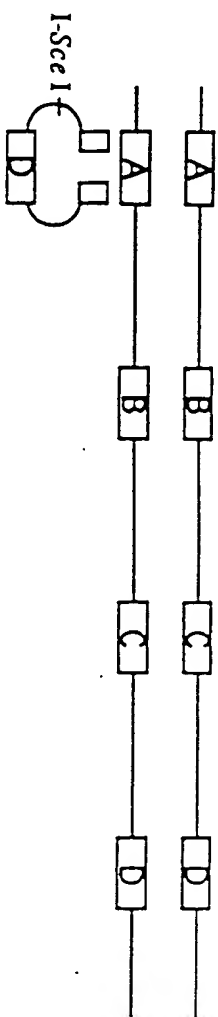
Recombination between the chromosome and plasmid



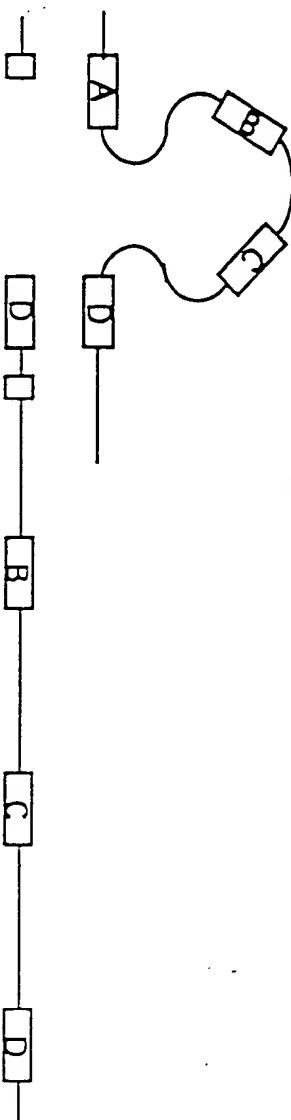
664000-001-0000

# DUPLICATION OF A LOCUS

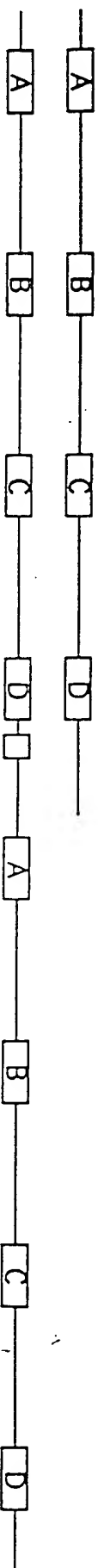
## 1 insertion of I-Sce I site by classical gene replacement



## 2 Specific cleavage by I-Sce I enzyme and repair of the break by homologous sequences



## 3 Duplication of the totality of the locus

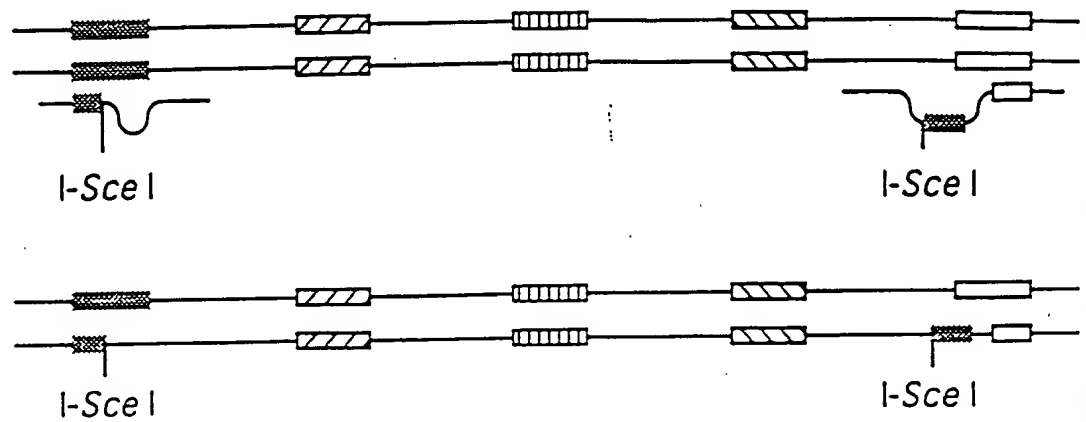


09244430.030493

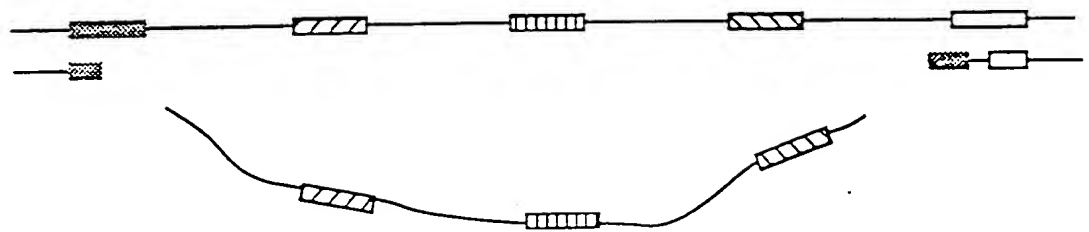


## DELETION OF A LOCUS

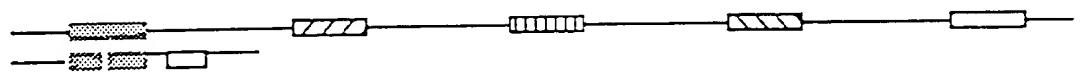
## 1 Insertion of two I-Sce I sites flanking the locus



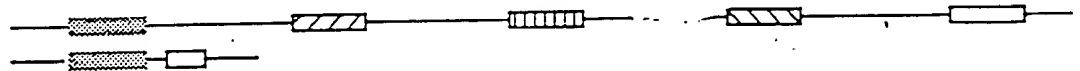
## 2 Expression of the enzyme and cleavage



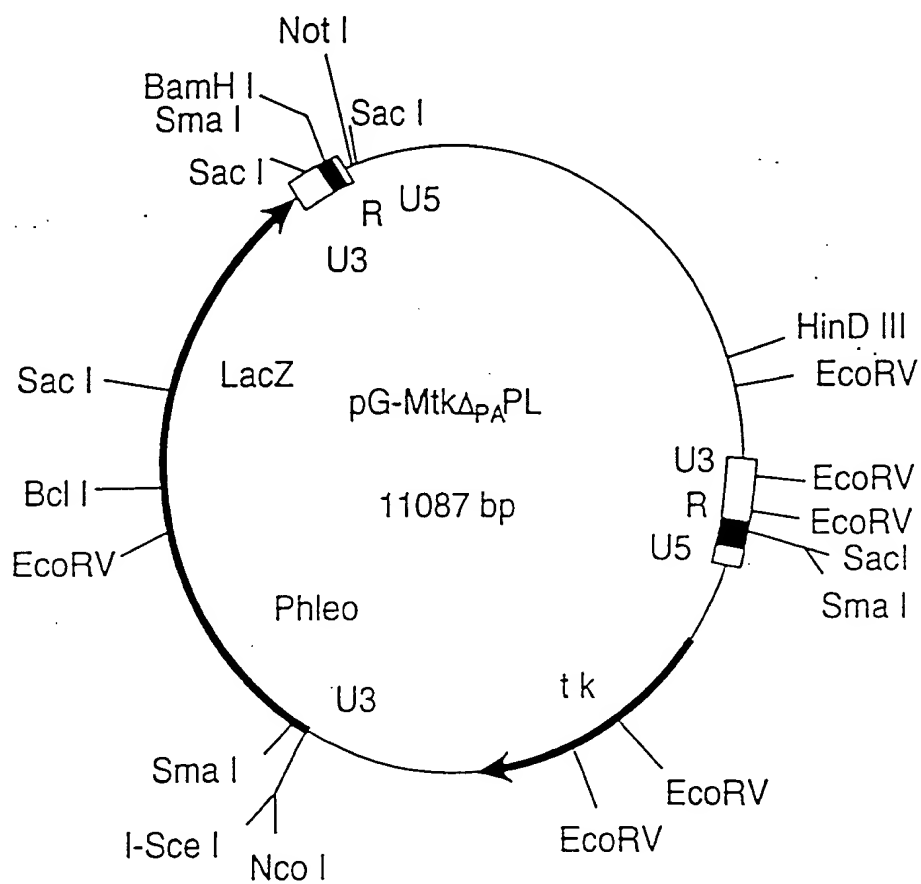
## 3 Recombination between the two ends



## 4 deletion of the locus



607000 "DEFINITION"



**FIG. 33**